Problem 1

a) According to the note

$$\mathbf{q(\omega)} \propto exp\{E_{q(\alpha_1,\ldots,\alpha_d,\lambda)}[lnp(y|x,\omega,\lambda) + lnp(\alpha_1,\ldots,\alpha_d) + lnp(\omega|\alpha_1,\ldots,\alpha_d) + \ln p(\lambda)]\}$$
$$\propto \exp\{\textstyle\sum_{i=1}^{N} E_{q(\lambda)}[\ lnp(y_i|x_i,\omega,\lambda)\big] + E_{q(\alpha_1,\ldots,\alpha_d)}[lnp(\omega|\alpha_1,\ldots,\alpha_d)]\}$$
$$\propto [\textstyle\prod_{i=1}^{N} e^{\frac{1}{2}E[\lambda]} e^{-\left(\frac{E_{q(\lambda)}[\lambda]}{2}\right)(y_i - x_i^T\omega)^2}] e^{-\frac{1}{2}\sum_{k=1}^{d} E_{q(\alpha_k)}[\alpha_k]\omega_k^2}$$

Therefore $q(\omega) = Normal(\omega|\mu', \Sigma')$

$$\mu' = \Sigma'(E_{q(\lambda)}[\lambda] \textstyle\sum_{i=1}^{N} y_i x_i)$$
$$\Sigma' = (diag\big(E_{q(\alpha_1)}[\alpha_1],\ldots,E_{q(\alpha_d)}[\alpha_d]\big) + E_{q(\lambda)}[\lambda] \textstyle\sum_{i=1}^{N} x_i x_i^T)^{-1}$$

$$\mathbf{q(\lambda)} \propto exp\{E_{q(\alpha_1,\ldots,\alpha_d,\omega)}[lnp(y|x,\omega,\lambda) + lnp(\alpha_1,\ldots,\alpha_d) + lnp(\omega|\alpha_1,\ldots,\alpha_d) + \ln p(\lambda)]\}$$
$$\propto \exp\{\textstyle\sum_{i=1}^{N} E_{q(\omega)}[\ lnp(y_i|x_i,\omega,\lambda)]\}p(\lambda)$$
$$\propto [\textstyle\prod_{i=1}^{N} \lambda^{\frac{1}{2}} e^{-\frac{\lambda}{2}E_{q(\omega)}[(y_i - x_i^T\omega)^2]}]\lambda^{e_0-1}e^{-f_0\lambda}$$

Therefore $q(\lambda) = Gamma(\lambda|e', f')$

$$e' = e_0 + \frac{N}{2}\ ,\quad f' = f_0 + \frac{1}{2}\textstyle\sum_{i=1}^{N} E_{q(\omega)}[\big(y_i - x_i^T\omega\big)^2]$$

For k=1,…,d

$$\mathbf{q(\alpha_k)} \propto exp\{E_{q(\omega,\lambda)\prod_{j\neq k}^{d} q(\alpha_j)}[lnp(y|x,\omega,\lambda) + lnp(\alpha_1,\ldots,\alpha_d) + lnp(\omega|\alpha_1,\ldots,\alpha_d) + \ln p(\lambda)]\}$$
$$\propto \exp\{lnp(\alpha_k) - \frac{1}{2}(-ln\alpha_k + E_{q(\omega)\prod_{j\neq k}^{d} q(\alpha_j)}\big[\omega^T diag\big(\alpha_1,\ldots,\alpha_d\big)\omega\big])\}$$
$$\propto e^{-\frac{1}{2}(-ln\alpha_k + E_{q(\omega)\prod_{j\neq k}^{d} q(\alpha_j)}[\omega^T diag(\alpha_1,\ldots,\alpha_d)\omega])} \alpha_k^{a_0-1} e^{-b_0\alpha_k}$$

Therefore $q(\alpha_k) = Gamma(\alpha_k|a', b_k')$

$$b_k' = b_0 + \frac{1}{2}(\Sigma'(k,k) + \mu'(k)^2)$$
$$a' = a_0 + \frac{1}{2}$$

b)

Inputs: Data and definitions $q(\omega) = Normal(\omega|\mu', \Sigma')$, $q(\lambda) = Gamma(\lambda|e_0', f_0')$ and $q(\alpha_k) = Gamma(\alpha_k|a_0, b')$

Outputs: Values for $\mu', \Sigma', e', f', a'$ and b_k'

1. initialize μ_0', Σ_0', e_0', f_0', a_0', b_{k0}' in some way
2. For iteration t = 1,…,T

   Update q(λ) by setting

$$e_t' = e_0 + \frac{N}{2}\ ,\quad f_t' = f_0 + \frac{1}{2}(\textstyle\sum_{i=1}^{N}\big(y_i - x_i^T\mu_{t-1}'\big)^2 + x_i^T\Sigma_{t-1}' x_i)$$

Update q(ω) by setting

$$\mu_t' = \Sigma_t'(\frac{e_t'}{f_t'}\sum_{i=1}^{N} y_i x_i)$$

$$\Sigma_t' = (diag\left(\frac{a_{t-1}'}{b_{1(t-1)}'}, \dots, \frac{a_{t-1}'}{b_{d(t-1)}'}\right) + \frac{e_t'}{f_t'}\sum_{i=1}^{N} x_i x_i^T)^{-1}$$

Updating q(α) by setting

$$b_{tk}' = b_0 + \frac{1}{2}(\Sigma_t'(k,k) + \mu_t'(k)^2)$$

$$a_t' = a_0 + \frac{1}{2}$$

Evaluate $\mathcal{L}(e_t', f_t', \mu_t', \Sigma_t', a_t', b_t')$ to access convergence

c)

$$\mathcal{L}(e_t', f_t', \mu_t', \Sigma_t', a_t', b_t') = E_q[lnp(y, \omega, \lambda, \alpha|x)] - E_q[lnq(\omega)] - E_q[lnq(\lambda)] - E_q[lnq(\alpha)]$$

$$= \frac{N}{2}(\psi(e_t') - ln2\pi f_t') - \frac{e_t'}{f_t'}(f_t' - f_0) +$$

$$e_0 lnf_0 - ln\tau(e_0) - \frac{f_0 e_t'}{f_t'} + (e_0 - 1)(\psi(e_t') - \ln f_t') +$$

$$\frac{d}{2}(\psi(a_t') - ln2\pi) - \sum_{k=1}^{d}\left(\frac{a_t'}{b_{kt}'}(b_{kt}' - b_0) + \frac{1}{2}lnb_{kt}'\right) +$$

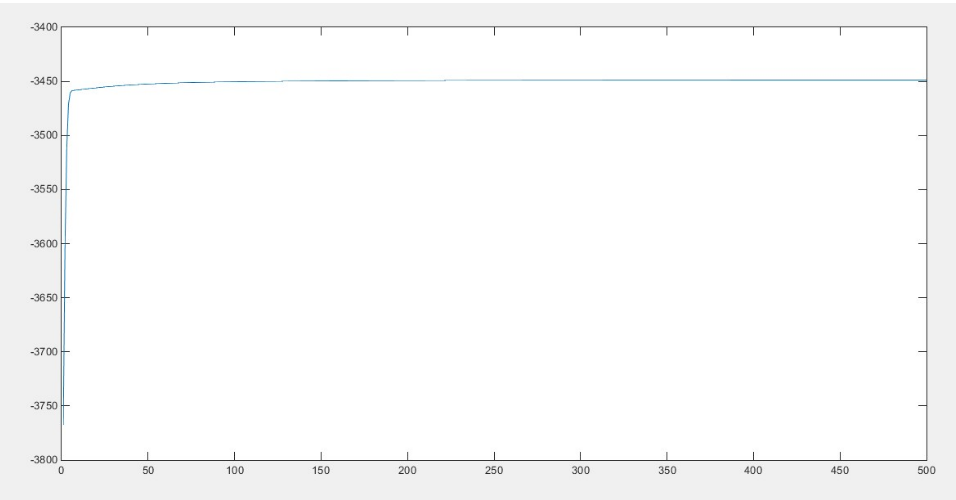
$$\mathrm{d}\big(a_0 lnb_0 + (a_0 - 1)\psi(a_t') - ln\tau(a_0)\big) - \sum_{k=1}^{d}\left(\frac{b_0 a_t'}{b_{kt}'} + (a_0 - 1)lnb_{kt}'\right) +$$

$$\frac{1}{2}\ln\left|2\pi e\Sigma_t'\right| - \ln f_t' + ln\tau(e_t') + e_t' - (e_t' - 1)\psi(e_t') -$$

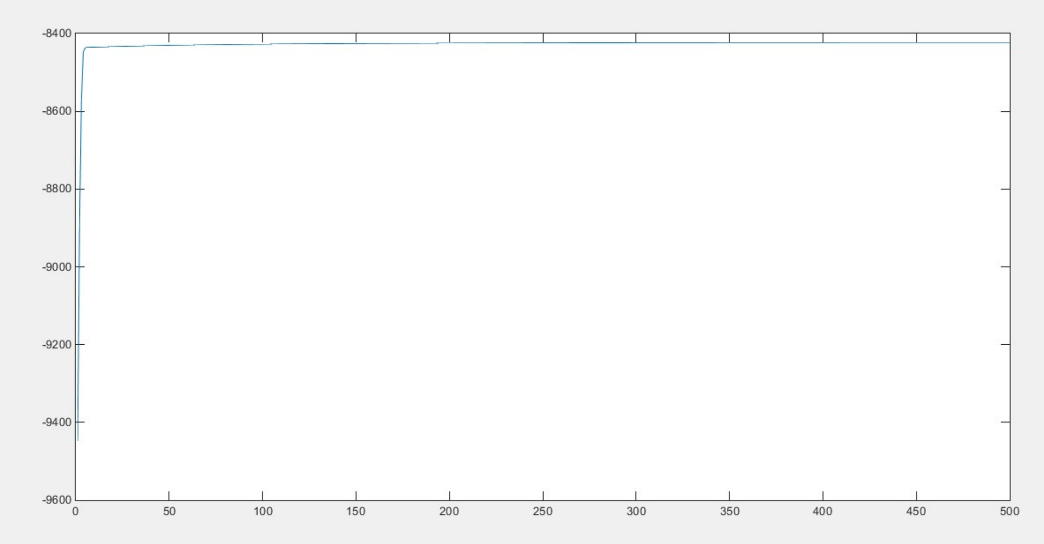
$$\mathrm{d}\big((a_t' - 1)\psi(a_t') - ln\tau(a_t') - a_t'\big) - \sum_{k=1}^{d} lnb_{kt}'$$

Problem 2

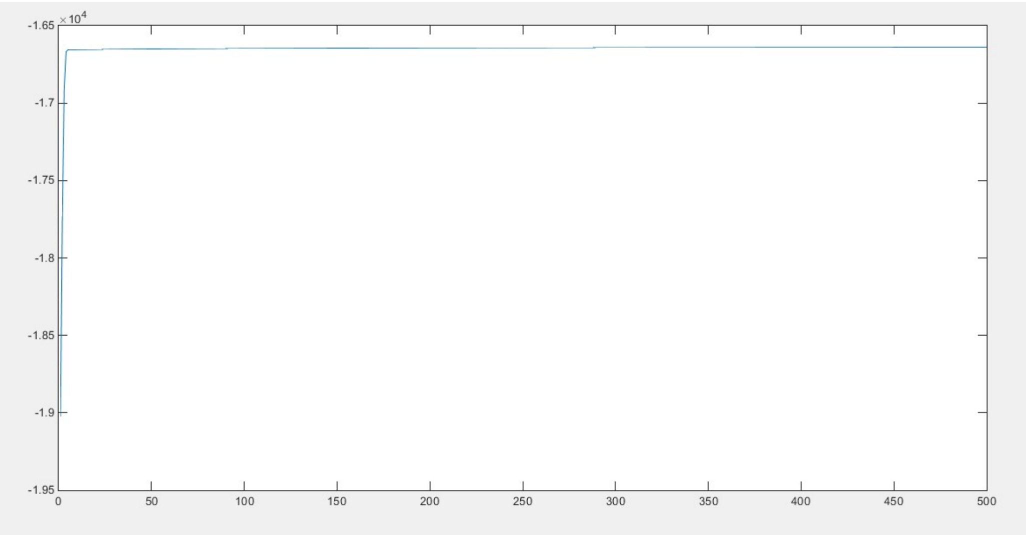
a) For the first dataset



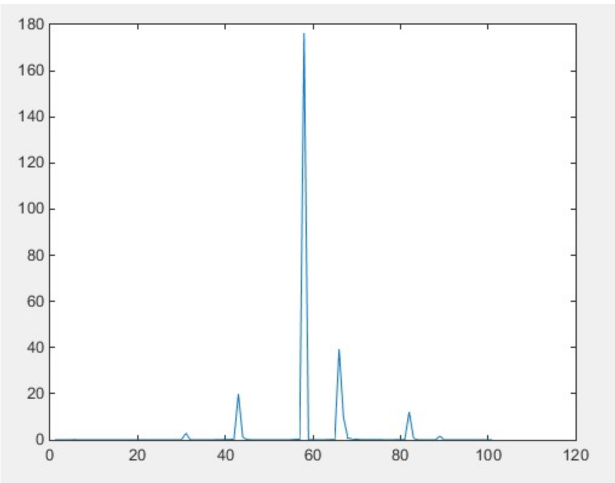
For the second dataset



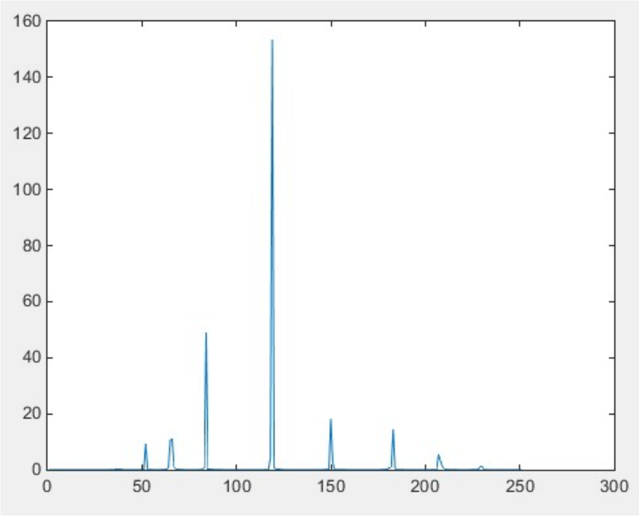
For the third dataset



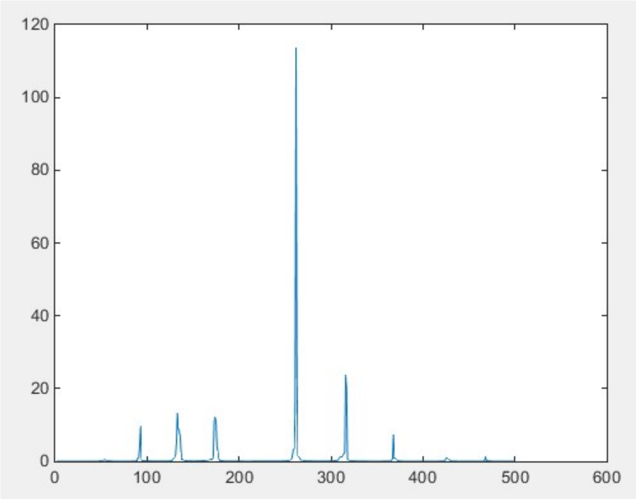
b) For the first dataset



For the second dataset



For the third dataset



c) For the first dataset

```
1/Eq[λ]=1.07983e+00
```

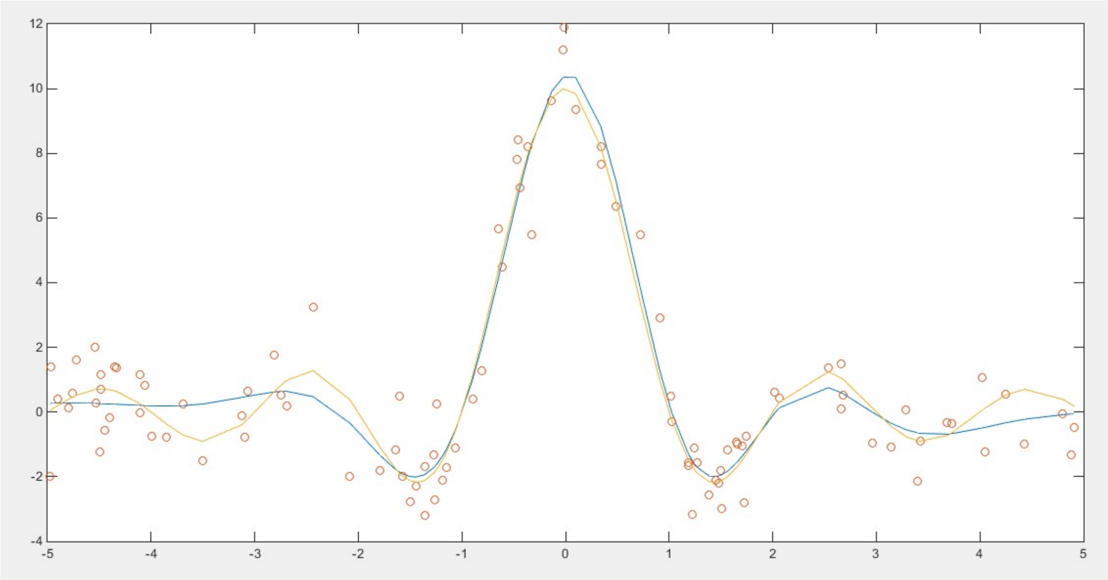
For the second dataset

```
1/Eq[λ]=8.99466e-01
```

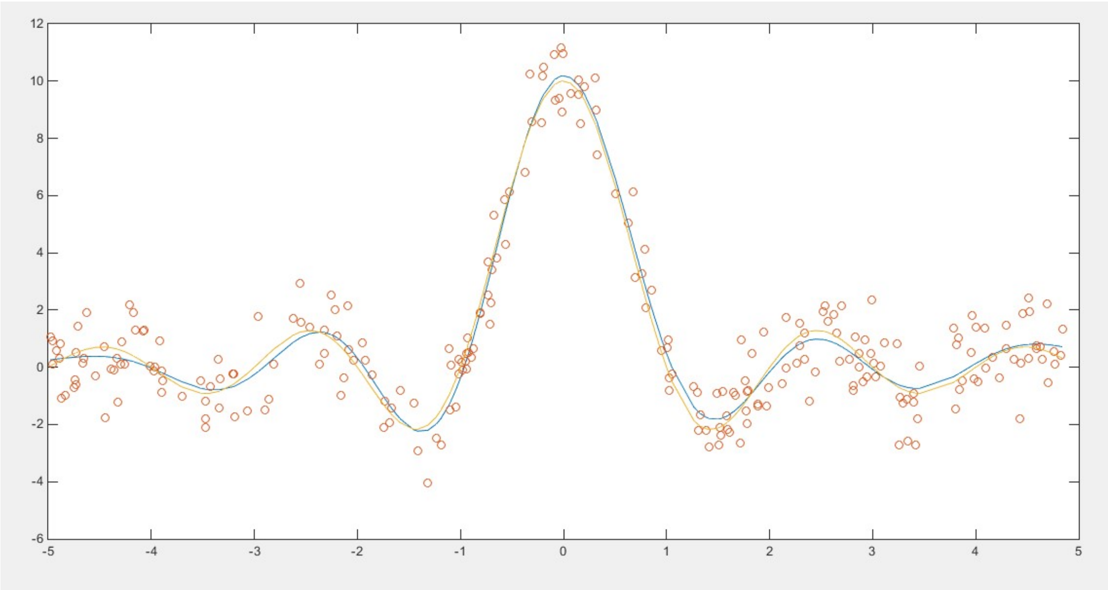
For the third dataset

```
1/Eq[λ]=9.78142e-01
```

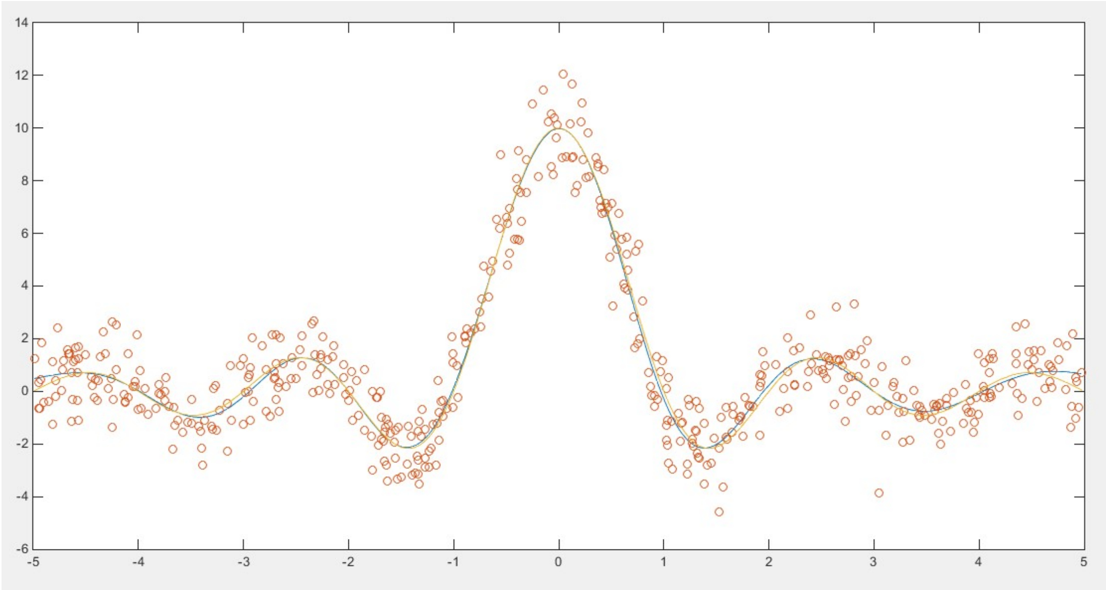
d) For the first dataset



for the second dataset



for the third dataset


14
12
10
8
6
4
2
0
-2
-4
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5